

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

Table of Contents

2000 VSD Data Dictionary	1
Constant File (Youth and Adult cohorts)	1
Enroll File (Youth and Adult cohorts)	4
Vaccine File (Youth and Adult cohorts)	7
Inpt (Inpatient) File (Youth and Adult cohorts)	18
Outpt (Outpatient) File (Youth and Adult [specified dxcodes only - see attached Adultoutptcodes.xls MS Excel file] cohorts)	22
Hmobirth (HMO Birth) File (Youth cohort)	26
Mort (Mortality) File (Youth and Adult cohort combined)	30
Address File (Youth cohort)	32
Geocode File (Youth cohort)	34
 Attachments	 36
I - Non-Standard ICD-9 Codes	36
II - Numeric Length Specifications	37
III - ICD-9 Code Conventions	38
IV - Facility Codes	40
V - Startup Dates for NCK Facilities	43
VI - Department Codes	see attached Deptcode.xls MS Excel file

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 1: CONSTANT File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	Note 1
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
	SAMPLE	N(4)	Z4.	Random, QA/QC Analysis Number	Uniform random number (0000-9999)	Note 2 (See 'Numeric Lengths' attachment.)
	BRTHDATE	N(4)	mmddyy10.	Date of Birth	SAS date . = Unknown	Notes 3a-3c
	SEX	C(1)			'F' = Female 'M' = Male ' ' = Unknown	

TABLE 1: CONSTANT File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	MC2000	N(3)	binary12.	Monthly Medicaid Status	<p>12 months of the year are each represented by one bit. The most significant bit corresponds to January, and the least significant bit corresponds to December.</p> <p>0 bit = not on Medicaid, unk, or missing 1 bit = Medicaid coverage</p> <p>Examples of stored SAS numeric value displayed with binary12. format:</p> <p>No Medicaid coverage (MC2000=0): 000000000000</p> <p>Medicaid coverage in January only (MC2000=2048): 100000000000</p> <p>Medicaid coverage in December only (MC2000=1): 000000000001</p> <p>Medicaid coverage every month of 2000 (MC2000=4095): 111111111111</p>	

Notes:

1. ID may be any combination of alphanumerics so long as they are unique among study participants and remain the same for each participant from cycle to cycle. Do not include leading, embedded or trailing blanks or special characters.
2. $SAMPLE = FLOOR(10000 * X)$, where X is the output of the SAS RANUNI function. All values returned should be zero-filled and right justified.
- 3a. The *Youth* cohort consists of:
 - (1) GHC - people born between March 2, 1973 and December 31, 2000, and having at least one day of VSD eligibility while less than 18 years old.

TABLE 1: CONSTANT File (contd.)

Notes (contd.):

- (2) KPC - people born between January 2, 1977 and December 31, 2000, and having at least one day of VSD eligibility while less than 18 years old.
- (3) NCK - people born between April 30, 1975 and December 31, 2000, and having at least one day of VSD eligibility while less than 18 years old.
- (4) HAP & NWK - people born between January 2, 1973 and December 31, 2000, and having at least one day of VSD eligibility while less than 18 years old.
- (5) SCK - people born between October 2, 1974 and December 31, 2000, and having at least one day of VSD eligibility while less than 18 years old.
- 3b. The **Adult** cohort consists of:
GHC, NCK, & NWK - people born on or before December 31, 1982, and having at least one day of VSD eligibility while aged 18 and older.
- 3c. SCK - If the actual day of birth is unknown, a *default value of the first of the month* is assigned within the HMO's membership database.

General file description:

This file has a single record for each VSD study member who has ever been enrolled in the study. Each STUDYID in the CONSTANT file should have at least one ENROLL record.

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 2: ENROLL File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	Note 1
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
Y	STRTDATE	N(4)	mmddyy10.	LLDB Start Date	Beginning of study <u>or</u> study enrollment date (if after beginning of study) - SAS date	Notes 2a-2c
	STOPDATE	N(4)	mmddyy10.	LLDB Stop Date	Drop date <u>or</u> last day of cycle's collection period - SAS Date	Notes 3a-3e

Notes:

1. ID may be any combination of alphanumerics so long as they are unique among study participants and remain the same for each participant from cycle to cycle. Do not include leading, embedded or trailing blanks or special characters.
- 2a. If enrolled in HMO before VSD study began, STRTDATE is HMO's VSD study start date:
Youth cohort: GHC & NCK* = 3/1/1991, HAP & NWK = 1/1/1991, SCK = 10/1/1992, and KPC = 1/1/1995.
 *NCK's child data (age < 7): STRTDATE is the day facility began collecting immunization data.
 NCK's adolescent data (aged 7-17): STRTDATE = 4/19/1993.
Adult cohort: GHC, NCK, & NWK = 1/1/1995.

TABLE 2: ENROLL File (contd.)

Notes (contd.):

- 2b. If enrolled (not birth) in HMO after VSD study began, STRTDATE is:
- (1) GHC & HAP - actual enrollment day.
 - (2) KPC - first of the next month.
 - (3) NCK - 3/1/91 - 4/18/1993:
 - (a) enrolled before facility was on-line - the day facility began collection immunization data.
 - (b) enrolled after facility was on-line - first date of immunization or well-care visit in VSD study clinic.
 - NCK - 4/19/1993 through present: after two months of contiguous membership, the 15th day of the 1st month of membership.
 - (4) NWK - first day of enrollment month.
 - (5) SCK - actual enrollment day.
- 2c. If birth in HMO after VSD study began, STRTDATE is:
- (1) GHC - actual enrollment day. If enrollment day < DOB, then STRTDATE=DOB.
 - (2) HAP - if born and enrolled within 60 days in HAP, then STRTDATE=DOB.
 - (3) KPC - STRTDATE=DOB.
 - (4) NCK - 3/1/1991 - 4/18/1993: first date of immunization or visit in on-line VSD study clinic.
4/19/1993 through present: DOB. If conflicting DOBs exist, the earliest available date of birth is used to populate STRTDATE. Note: For some STUDYIDs, this may cause enrollment STRTDATEs to occur prior to DOB. These DOBs are usually corrected over time.
 - (5) NWK - DOB if enrollment day is within 93 days of DOB and mother was HMO eligible.
 - (6) SCK - DOB if enrollment day is within 31 days of DOB.
- 3a. If active in HMO at end of cycle's collection period, STOPDATE is the last day of the cycle (e.g., 2000 [C9] STOPDATE: 12/31/2000).
- 3b. If transfer (NCK) - date participant has event at a non-VSD study clinic.
- 3c. If terminated prior to last day of the cycle:
- (1) GHC, HAP, & SCK - actual HMO termination date.
 - (2) NCK - 3/1/1991 - 4/18/1993: last day of termination month.
4/19/1993- present: after two months of contiguous non-membership, STOPDATE is set to the last day of the month prior to the two-month non-enrollment period.
 - (2) KPC & NWK - last day of termination month.
- 3d. If aged out of *Youth* cohort, STOPDATE is:
- (1) All HMOs: day before the eighteenth birthday.
- 3e. If deceased, STOPDATE is the definite or approximate date of death. If an approximate date of death is used, it will be updated as more accurate data becomes available.

General file description:

This file has a single record for each HMO enrollment episode that a VSD study member has during the study collection period. Enrollment periods should not overlap and consecutive enrollment periods should be combined into one enrollment period. Each STUDYID in the ENROLL file should have one record in the CONSTANT file. **Note:** Data outside a VSD study member's enrollment intervals should not be considered complete

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 3: VACCINE File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
Y	VACDATE	N(4)	mmddyy10.	Date of Vaccine Administration	SAS date . = Unknown	Notes 1a & 1b (See 'Numeric Lengths' attachment.)

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	VAC	C(3)		Vaccine Administered	'54' = Adenovirus, type 4 '55' = Adenovirus, type 7 '82' = Adenovirus, NOS '24' = Anthrax '19' = BCG (Bacillus Calmette-Guerin) '27' = Botulinum antitoxin '26' = Cholera '29' = CMVIG (cytomegalovirus IG, intravenous) '56' = Dengue fever '12' = Diphtheria antitoxin '28' = DT (peds) (diphtheria/tetanus) '20' = DTaP (DTacellP) '50' = DTaP-Hib (DTacellP w/Hib) 'DHB' = DtaP-HepB (DTacellP w/HepB) '01' = DTP (diphtheria/tetanus/pertussis) '22' = DTP-Hib (DTP w/Hib) '102' = DTP-Hib-HepB '57' = Hantavirus '52' = Hep A, adult '83' = Hep A, ped/adol, 2 dose '84' = Hep A, ped/adol, 3 dose '31' = Hep A, pediatric, NOS '85' = Hep A, NOS '30' = HBIG (hepatitis B IG) '08' = Hep B, adolescent or pediatric '42' = Hep B, adolescent/high risk infant '43' = Hep B, adult '44' = Hep B, dialysis '45' = Hep B, NOS '58' = Hep C	Notes 2 & 3. If an additional VAC value is needed, please contact CDC. Left justify New vaccine codes are in bold .

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	VAC (contd)	C(3)		Vaccine Administered	'59' = Hep E '60' = Herpes simplex 2 '46' = Hib (PRP-D) '47' = Hib (HbOC) '48' = Hib (PRP-T) '49' = Hib (PRP-OMP) '17' = Hib, NOS 'HBP' = Hib polysaccharide '51' = Hib-Hep B '61' = HIV (human immunodeficiency) '62' = HPV (human papilloma) '86' = IG (intramuscular) '87' = IGIV (intravenous) '14' = IG, NOS '15' = Influenza, split '16' = Influenza, whole '88' = Influenza, NOS '10' = IPV (poliovirus, inactivated) '02' = OPV (poliovirus, live, oral) '89' = Polio, NOS '39' = Japanese encephalitis '63' = Junin virus '64' = Leishmaniasis '65' = Leprosy '66' = Lyme disease '03' = MMR (measles/mumps/rubella) '94' = MMR-V (MMR with varicella) '04' = M/R (measles/rubella) '67' = Malaria '05' = Measles	

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	VAC (contd)	C(3)		Vaccine Administered	'MM' = Measles/Mumps 'RU' = Measles/Rubeola (Attenuvax) '68' = Melanoma '32' = Meningococcal '103' = Meningococcal C conjugate '07' = Mumps '69' = Parainfluenza-3 '11' = Pertussis '23' = Plague '33' = Pneumococcal '100' = Pneumococcal conjugate vaccine, polyvalent '70' = Q fever '18' = Rabies, intramuscular injection '40' = Rabies, intradermal injection '90' = Rabies, NOS '72' = Rheumatic fever '73' = Rift Valley Fever '34' = RIG (rabies IG) '74' = Rotavirus '71' = RSV-IGIV (respiratory syncytial virus IG, intravenous) '93' = RSV-MAb (respiratory syncytial virus monoclonal antibody [palivizumab], intramuscular) '06' = Rubella '38' = Rubella/Mumps '75' = Smallpox '76' = Staphylococcus bacterio lysate '09' = Td (adult) (tetanus/diphtheria) '35' = Tetanus toxoid '77' = Tick-borne encephalitis	

TABLE 3: VACCINE File (contd.)

[illegible]

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	SITE	C(2)		Site of Vaccine Introduction	'IN' = Intranasal 'IV' = Intravenous 'LA' = Left arm 'LD' = Left deltoid 'LG' = Left upper outside quad 'LL' = Left lateral thigh 'LT' = Left anterior thigh 'LU' = Left thigh, unknown position 'LV' = Left gluteus 'PO' = Oral 'RA' = Right arm 'RD' = Right deltoid 'RG' = Right upper outside quad 'RL' = Right lateral thigh 'RT' = Right anterior thigh 'RU' = Right thigh, unknown position 'RV' = Right gluteus 'UA' = Unknown arm, left or right 'UD' = Unknown deltoid, left or right 'UT' = Unknown thigh, left or right 'U ' = Unavailable ' ' = Unknown	

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	MFR	C(3)		Vaccine Manufacturer	'AB ' = Abbott Laboratories 'AD ' = Adams Laboratories 'ALP' = Alpha Therapeutic Corporation 'AR ' = <i>Armour</i> 'AVB' = Aventis Behring L.L.C. (formerly Centeon L.L.C., includes Armour) 'AVI' = Aviron 'BA ' = <i>Baxter Healthcare Corporation</i> 'BAH' = Baxter Healthcare Corporation (includes Hyland Immuno, Immuno International AG, & North American Vaccine, Inc) 'BAY' = Bayer Corporation 'BP ' = <i>Berna Products</i> 'BPC' = Berna Products Corporation 'BCP' = BIOChem Pharmaceutical 'CEN' = <i>Centeon L.L.C.</i> 'CHI' = Chiron Corporation 'CMP' = Celltech Medeva Pharmaceuticals 'CON' = <i>Connaught</i> 'EVN' = <i>Evans Medical Limited</i> 'GRE' = Greer Laboratories, Inc. 'IAG' = <i>Immuno International AG</i> 'IM ' = <i>Merieux</i> 'IUS' = Immuno-U.S., Inc. 'JPN' = The Research Foundation for Microbial Diseases of Osaka University (BIKEN) 'KGC' = Korea Green Cross Corporation	Notes 7 and 8. If an additional MFR value is needed, please contact CDC. Left justify. New manufacturer codes are in bold . MFRs which have been through mergers, etc. are in <i>italics</i> .

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	MFR (contd)	C(3)		Vaccine manufacturer	<p><i>'LED'</i> = Lederle</p> <p><i>'MA '</i> = Massachusetts Public Health Biologic Laboratories</p> <p>'MBL' = Massachusetts Biologic Laboratories (formerly Massachusetts Public Health Biologic Laboratories)</p> <p><i>'MED'</i> = MedImmune, Inc.</p> <p><i>'MIL'</i> = Miles</p> <p><i>'MIP'</i> = Bioport Corporation (formerly Michigan Biologic Products Institute)</p> <p><i>'MSD'</i> = Merck & Co., Inc.</p> <p><i>'NAB'</i> = NABI (formerly North American Biologicals, Inc.)</p> <p><i>'VS1'</i> = National Institute/Agency International Development</p> <p><i>'VS2'</i> = National Institute of Health</p> <p><i>'NYB'</i> = New York Blood Center</p> <p><i>'NAV'</i> = North American Vaccine, Inc.</p> <p><i>'NOV'</i> = Novartis Pharmaceutical Corp.</p> <p><i>'OTC'</i> = Organon Teknika Corporation</p> <p><i>'ORT'</i> = Ortho-Clinical Diagnostic Systems, Inc. (formerly Ortho Diagnostic Systems, Inc.)</p> <p><i>'PD '</i> = Parke-Davis (formerly Parke- Davis)</p>	

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	MFR (contd)	C(3)		Vaccine manufacturer	<p>‘PMC’ = Aventis Pasteur Inc. (formerly Pasteur Merieux Connaught; includes Connaught Laboratories and Pasteur Merieux)</p> <p>‘PRX’ = <i>Praxis Biologics</i></p> <p>‘PWJ’ = Powerject Pharmaceuticals (includes Celltech Medeva Vaccines and Evans Medical Limited)</p> <p>‘SCL’ = Sclavo, Inc.</p> <p>‘VS3’ = SmithKline French</p> <p>‘SI ’ = <i>Swiss Serum and Vaccine Inst.</i></p> <p>‘SKB’ = GlaxoSmithKline (formerly SmithKline Beecham; includes SmithKline Beecham & Glaxo Wellcome)</p> <p>‘VS4’ = UCLA (SCK exp)</p> <p>‘USA’ = United States Army Medical Research & Materiel Command</p> <p>‘WA ’ = <i>Wyeth-Ayerst</i></p> <p>‘WAL’ = Wyeth Ayerst (includes Wyeth-Lederle Vaccines and Pediatrics, Wyeth Laboratories, Lederle Laboratories, & Praxis Biologics)</p> <p>‘OTH’ = Other</p> <p>‘UNK’ = Unknown</p>	

TABLE 3: VACCINE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	LOT	C(15)		Vaccine Lot Number	alphanumerics, left-justified no spaces, slashes, or dashes ' ' = Unknown	Notes 9 and 10.
	VACSOURC	C(1)		Location Where Vaccine Was Administered	'A' = Vaccinated at a HMO affiliate or partnership facility 'B' = Vaccinated at HMO before immunization system automated at clinic 'T' = Vaccinated within the HMO 'O' = Vaccinated outside the HMO 'U' = Unavailable ' ' = Unknown	

Notes:

- 1a. *Youth* cohort: All *Youth* Studyids' VACDATES which fall on or between the VSD *Youth* Study Start Date and one (1) day prior to a participant's 18th birthday, or the Cycle's Stop Date (whichever is earlier) will be submitted.
Note: Data outside of enrollment intervals should **not** be considered complete for *Youth* cohort.
- 1b. *Adult* cohort: All *Adult* Studyids' VACDATES which fall on or between the VSD *Adult* Study Start Date and the Cycle's Stop Date will be submitted, except data prior to the Studyid's 18th birthday.
Note: Data outside of enrollment intervals should **not** be considered complete for *Adult* cohort.
2. Health Level 7 (HL7) Standard Code Set CVX -- Vaccines Administered is maintained by the Centers for Disease Control and Prevention. For further information on the CVX code set, see <http://www.cdc.gov/nip/registry> . New vaccine codes and names are in bold.
3. NOS = not otherwise specified. Avoid using NOS codes except to record historical records that lack the indicated specificity or in cases where charts/automated systems do not capture the indicated specificity.
4. This vaccine value, 'Unn', represents a combination vaccine where the exact vaccine combination is unknown, but one or more of

TABLE 3: VACCINE File (contd.)

- the components are known. For example: 'U01' would be a combination of D, T and P along with other, unknown components. These occur mostly because of historical and/or non-HMO vaccines with illegible handwriting.
5. The vaccine values, 'Xnn', 'Ynn', and 'Znn', represent experimental vaccines being tested by NCK, SCK, & KPC, respectively. For each unique experimental vaccine, 'nn' should be replaced by a unique numeric value (e.g., X01, Y01).
 6. Records with missing VACs should be included. These records should be validated to ensure that the vaccine is indeed missing and can not be obtained through other avenues.
 7. Health Level 7 (HL7) Standard Code Set MVX – Manufacturers of Vaccines is maintained by the Centers for Disease Control and Prevention. For further information on the MVX code set, see <http://www.cdc.gov/NIP/Registry/cvx.pdf> and <http://www.cdc.gov/NIP/Registry/mvx.pdf>. Vaccine names and codes that have been through mergers, etc., are in italics. New manufacturer codes and names are in bold.
 8. The MFR variable should be coded with the HL7 MVX value who manufactured the vaccine. For example, vaccines administered prior to Praxis Biologics merger with Wyeth-Ayerst, should be coded as 'PRX'.
 9. If combination vaccine, enter the container's lot number.
 10. Format LOT variable as follows: alphanumerics, left-justified with no spaces, slashes, or dashes between characters, and with no dose or dose-unit information attached (e.g., EJN123).

General file description:

This file contains a single record for each vaccine administered during the study collection period. Records with missing VACs should be included if documentation exists that a vaccine was administered on a specific date, but the vaccine administered is unknown. Each STUDYID in the VACCINE file should have one record in the CONSTANT file and one or more records in the ENROLL file. **Note:** Data outside of enrollment intervals should not be considered complete.

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 4: INPT (Inpatient) File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
Y	CAREDATE	N(4)	mmddyy10.	Date of Admission	SAS date . = Unknown	Notes 1a-1b (See 'Numeric Lengths' attachment.)
	DURATION	N(3)		Length of Stay	NN = Length of stay .U = Unavailable . = Unknown	Note 2

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	DXCODE	C(6)		ICD-9 Code	ICD-9 code (No decimal) ' ' = Unknown	Notes 3a-3b (See 'ICD-9 Code Convention' attachment.)
Y	DXTYPE	C(2)		Setting	'BP' = Birth hosp primary DX 'BS' = Birth hosp secondary DX 'BN' = Birth hosp, DX designation N/A 'HP' = Hosp primary DX 'HS' = Hosp secondary DX 'HN' = Hosp, DX designation N/A ' ' = Unknown	Notes 4a-4b Notes 4a-4b Notes 4a-4b Note 4b Note 5
Y	FACILITY	C(3)		Facility ID Code	Specific facility code where outcome occurred 'N' = Not applicable 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	Left justify (See 'Facility Codes' attachment)
Y	LOCATION	C(1)		In/Outside HMO	'A' = HMO affiliated or partnership facility 'I' = Inside HMO 'O' = Outside HMO 'R' = Referral 'U' = Unavailable ' ' = Unknown	

TABLE 4: INPT File (contd.)

Notes:

- 1a. **Youth cohort:** All *Youth* Studyids' CAREDATES which fall on or between the VSD Youth Study Start Date and one (1) day prior to a participant's 18th birthday, or the Cycle's Stop Date (whichever is earlier) will be submitted. Youth cohort hospital history (formerly a stand alone file, HOSPDXXH, for GHC, NCK, NWK and SCK) will be appended to the INPT (INPATIENT) file and extend back to 1/1/1984 for GHC, NCK, and NWK, 6/1/1984 for SCK, 1/1/1988 for KPC and 1/1/1990 for HAP.
Note: Data outside of enrollment intervals should **not** be considered complete for *Youth* cohort.
- 1b. **Adult cohort:** All *Adult* Studyids' CAREDATES which fall on or between the VSD Adult Study Start Date and the Cycle's Stop Date will be submitted. Adult cohort hospital history (formerly a stand alone file, HOSPDXXH, for GHC, NCK, and NWK) will be appended to the INPT (INPATIENT) file and extend back to 1/1/1990 for GHC, NCK, and NWK.
Note: Data outside of enrollment intervals should **not** be considered complete for *Adult* cohort.
- 2. DURATION = Hospital discharge date - admission date. DURATION will be zero for hospitalizations with the same admission and discharge date.
- 3a. Codes should be left justified and not contain decimals. Codes below 100.nnn should have leading zeroes. See 'ICD-9 Code Convention' attachment.
- 3b. Records with missing DXCODEs should be included. These records should be validated to ensure that the diagnosis is indeed missing and can not be obtained through other avenues.
- 4a. DXTYPE = 'BP' should be assigned when the DXCODE has the first three digits of V30 - V39, and the 4th digit is 0. Associated records for the same admission would use 'BS'. If the infant was readmitted/examined on the same day as birth, or on a different day, then that re-admittance/examination should be coded differently; probably 'HP' or 'HS', but according to the actual setting.
- 4b. If Birth hospitalization or Hospitalization, and a distinction between primary/secondary diagnoses can not be made, code 'BN' or 'HN', respectively.
- 5. The Adult cohort contains **only** HP, HS, and HN DXTYPEs.

General file description:

This file contains a single record for each diagnosis assigned during a hospitalization during the study collection period. Records with missing DXCODEs should be included if documentation exists that an INPT encounter occurred on a specific date, but the diagnosis is missing. Hospital history prior to the study collection period is also included. If a VSD study member has two or more hospital visits occurring on the same day for the same dxcode, but occurring in different facilities or for different durations, each record should be submitted. Excluded (rule-out) diagnoses should not be included in this file. Each STUDYID in the INPT file should have one record in the CONSTANT file and one or more records in the ENROLL file.

Note: Data outside of enrollment intervals should not be considered complete.

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 5: OUTPT (Outpatient) File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	Notes 1a-1c
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
Y	CAREDATE	N(4)	mmddyy10.	Date of Visit	SAS date . = Unknown	Note 1b (See 'Numeric Lengths' attachment.)
Y	DEPT	C(3)		Department	See Attachment VI for coding	Left justify, please refer to Attachment VI for code descriptions

TABLE 5: OUTPT File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	DXCODE	C(6)		ICD-9 Code	ICD-9 code (No decimal) ' ' = Unknown	Notes 1a & 2 (See 'ICD-9 Code Convention' attachment.)
Y	DXTYPE	C(2)		Setting	'CP' = Clinic visit primary DX 'CS' = Clinic visit secondary DX 'CF' = Clinic follow-up visit DX 'CN' = Clinic visit, DX designation N/A 'EP' = Emergency room primary DX 'ES' = Emergency room secondary DX 'EF' = Emergency room follow-up visit DX 'EN' = Emergency room, DX designation N/A 'UP' = Urgent care visit primary DX 'US' = Urgent care visit secondary DX 'UF' = Urgent care follow-up visit DX 'UN' = Urgent care visit, DX designation N/A ' ' = Unknown	Note 3a Note 3b Note 3a Note 3b Note 3a Note 3b
Y	FACILITY	C(3)		Facility ID Code	Specific facility code where outcome occurred 'N' = Not applicable 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	Left justify (See 'Facility Codes' attachment)

TABLE 5: OUTPT File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	LOCATION	C(1)		In/Outside HMO	'A' = HMO affiliated or partnership facility 'I' = Inside HMO 'O' = Outside HMO 'R' = Referral 'U' = Unavailable ' ' = Unknown	
	RULEOUT	C(1)		Rule-out DX	'R' = Rule-out code ' ' = No rule-out code assigned 'U' = Unavailable	Note 4

Notes:

- 1a. The OUTPT file is submitted for the Youth and Adult cohorts; however, for the Adult cohort it is limited to *high risk Influenza, Autoimmune Disease, Chickenpox (052.*) and Herpes Zoster (053.*) ICD-9-CM codes*. (See Adult Outpt File DXCODEs document.)
- 1b. **Youth cohort:** All *Youth* Studyids' CAREDATES which fall on or between the VSD Youth Study Start Date and one (1) day prior to a participant's 18th birthday, or the Cycle's Stop Date (whichever is earlier) will be submitted.
Note: Data outside of enrollment intervals should **not** be considered complete for *Youth* cohort.
- 1c. **Adult cohort:** All *Adult* Studyids' CAREDATES which fall on or between the VSD Adult Study Start Date and the Cycle's Stop Date, and have the pre-specified ICD-9-CM codes, will be submitted.
Note: Data outside of enrollment intervals should **not** be considered complete for *Adult* cohort.
2. Codes should be left justified and not contain decimals. Codes below 100.nnn should have leading zeroes. See 'ICD-9 Code Convention' attachment.
- 3a. If ER, Urgent Care, or Clinic *'follow-up'* visit diagnoses, code 'EF', 'UF', or 'CF', respectively. This code is not available for all occurrences of follow-up visits, but will be utilized when possible.

TABLE 5: OUTPT File (contd.)

- 3b. If Emergency Room, Urgent Care, or Clinic visit and a distinction between primary/secondary diagnoses can not be made, code 'EN', 'UN', or 'CN', respectively.
- 4. Rule-out codes are used with ER, Urgent Care visits, and/or Clinic visits only. Rule-out indicates that the ICD-9 code is under investigation and the investigation is being done to 'rule-out' the ICD-9 diagnoses code.

General file description:

Youth: This file contains a single record for each diagnosis assigned during an emergency room, urgent care, or clinic visit during the study collection period. If a VSD study member has two or more visits occurring on the same day for the same dxcode, but occurring in different facilities or departments, each record should be submitted. Each STUDYID in the OUTPT file should have one record in the CONSTANT file and one or more records in the ENROLL file. **Note:** Data outside of enrollment intervals should not be considered complete.

Adult: This file contains a single record for pre-specified high-risk influenza, autoimmune disease, herpes zoster, and chicken pox diagnoses assigned during an emergency room, urgent care, or clinic visit during the study collection period. If a VSD study member has two or more visits occurring on the same day for the same dxcode, but occurring in different facilities or departments, each record should be submitted. Each STUDYID in the OUTPT file should have one record in the CONSTANT file and one or more records in the ENROLL file. **Note:** Data outside of enrollment intervals should not be considered complete.

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 6: HMOBIRTH File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	Notes 1 & 2
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
	MOMID	C(8)		Mother's VSD ID Number	Unique combination of alphanumerics	Note 3
	FACILITY	C(3)		Facility ID Code	Specific facility code where birth occurred 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	Left justify (See 'Facility Codes' attachment)
	GESTAGE	N(3)		Gestational Age	NN = Gestational Age in Weeks .U = Unavailable (HMO can not provide this variable) . = Unknown	

TABLE 6: HMOBIRTH File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	CHLDWGHT	N(4)		Child's Birth Weight	NNNN = Birth Weight in grams .U = Unavailable (HMO can not provide this variable) . = Unknown	
	APGAR1	C(2)		One Min Apgar	'00' - '10' = 1 minute APGAR score 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	
	APGAR5	C(2)		Five Min Apgar	'00' - '10' = 5 minute APGAR score 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	
	WHITE	C(1)			'1' = White	Note 4
	BLACK	C(1)			'1' = Black	Note 4
	ASIAN	C(1)			'1' = Asian	Note 4
	HAWAIIAN	C(1)		Hawaiian or Other Pacific Islander	'1' = Hawaiian or Other Pacific Islander	Note 4

TABLE 6: HMOBIRTH File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	NATAMER	C(1)		Native American or Aleutian	'1' = Native American or Aleutian	Note 4
	OTHER	C(1)		Other Race	'1' = Other Race	Note 4
	HISPANIC	C(1)			'1' = Hispanic	Note 4
	CHLDSTE	C(2)		Child's Birth State	Child's birth state FIPS code 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	
	CHLDCNTY	C(3)		Child's Birth County	Child's birth county FIPS code 'U' = Unavailable (HMO can not provide this variable) ' ' = Unknown	
	M_DOB	N(4)	mmddyy10.	Mother's Date of Birth	SAS date .U = Unavailable (HMO can not provide this variable) . = Unknown	

Notes:

TABLE 6: HMOBIRTH File (contd.)

1. ID may be any combination of alphanumerics so long as they are unique among study participants and remain the same for each participant from cycle to cycle. Do not include leading, embedded or trailing blanks or special characters.
2. This file consists of infants born in HMO's hospitals or contract facilities and are members of the VSD study population. Infants who are born on or after HMO's VSD study start date are included (GHC & NCK = 3/1/1991, HAP = 1/1/1991, KPC= 1/1/1995, NWK = 1/1/1991 and SCK = 10/1/1992). HMOs which match to state birth certificate files may have data on infants born prior to their VSD study start date as well as those born outside their HMO's hospitals or contract facilities.
3. For those HMOs who participate in the Adult cohort (GHC, NCK, and NWK), the birth mother's STUDYID is provided.
4. Complete each race/ethnicity category that applies.

General file description:

This file contains a single record for each Youth cohort member born in HMOs' hospitals or contract hospitals during the study collection period. For sites choosing to match to their respective state birth certificate files, data on infants born in non-HMO facilities or prior to the study collection period, may also be included. Each STUDYID in the HMOBIRTH file should have one record in the CONSTANT file and one or more records in the ENROLL file.

Vaccine Safety Datalink Project

TABLE 7: MORT (Mortality) File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	Note 1
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
	DXCODE	C(6)		Underlying Cause of Death	ICD-10 (No decimal) ' ' = Unknown	Notes 2a-2c
	DTHDATE	N(4)	mmddyy10.	Date of Death	SAS date ' ' = Unknown	
	APPROX	C(1)		Approximate Death Date	'Y' = Yes 'N' = No ' ' = Unknown	Note 3

Notes:

1. All known deaths should be submitted, regardless of date or age (i.e., even if they are after the STUDYID's last stop date, or after the Cycle's Stop Date). Ensure that the STUDYID also appears in either the Youth or Adult cohort's CONSTANT and ENROLLment files; if not, please update these files accordingly.

TABLE 7: MORT File (contd.)

- 2a. Codes should be left justified and not contain decimals. Codes below 100.nnn should have leading zeroes.
- 2b. Records with missing DXCODEs should be included. These records should be updated as the information becomes available.
- 2c. The source of the underlying cause of death (DXCODE) will be the state death file; otherwise it will be left blank. The missing DXCODEs will be updated as more accurate data becomes available. ICD-10 was implemented in the U.S. for mortality reporting as of 01/01/1999.
- 3. Indicates whether death date is definite death date or approximate. Approximate death dates will be updated as more accurate data becomes available.

General file description:

This file contains a single record for each known death among VSD youth and adult cohort members. If known, deaths that occur after the study collection period or after a VSD cohort members ages out, are also included. Each STUDYID in the MORT file should have one record in the youth and/or adult CONSTANT file and one or more records in the respective cohort(s) ENROLL file.

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 8: ADDRESS File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	STUDYID	C(8)		LLDB ID Number	Unique combination of alphanumerics	Notes 1a-1b
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
Y	POLLDATE	N(4)	mmddyy10.	Date Information Polled	SAS date	Notes 2a-2b (See 'Numeric Lengths' attachment.)
	ZIP	N(4)		Zip Code	5-digit zip code (nnnnn) .U = Unavailable . = Unknown	

TABLE 8: ADDRESS File (contd.)

Notes:

- 1a. The ADDRESS file is submitted for the *Youth* cohort only.
- 1b. ID may be any combination of alphanumerics so long as they are unique among study participants and remain the same for each participant from cycle to cycle. Do not include leading, embedded or trailing blanks or special characters.
- 2a. Each HMO will select two POLLDATEs: February 2000 and 2001 POLLDATEs.
A record will be created containing the address of the youth (responsible subscriber) on each POLLDATE.
- 2b. Youth should have a record for each POLLDATE when they were active. If they have no POLLDATE because they enrolled and quit between two POLLDATEs, or quit before their first POLLDATE, a single record should be created using the closest POLLDATE.

General file description:

This file contains Youth cohort members' United States Post Office zip code. The zip code may vary depending upon the POLLDATE. Each STUDYID in the ADDRESS file should have one record in the CONSTANT file and one or more records in the ENROLL file.

2000 VSD Data Dictionary

Vaccine Safety Datalink Project

TABLE 9: GEOCODE File

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
Y	GEOID	C(8)		Geocode ID	Scrambled alphanumeric characters	Notes 1a-1b Cannot be STUDYID
Y	CDCSITE	C(1)		HMO Site	'B' = Harvard Pilgrim 'C' = Northern California Kaiser 'D' = Kaiser Colorado 'O' = Northwest Kaiser 'W' = Group Health Cooperative 'S' = Southern California Kaiser	
Y	POLLDATE	N(4)	mmddyy10.	Date Information Polled	SAS date	Notes 2a-2b (See 'Numeric Lengths' attachment.)
	ADDRESS	C(33)			Street Address. If not available use P. O. Box 'U' = Unavailable ' ' = Unknown	

TABLE 9: GEOCODE File (contd.)

Key	Variable	Type (Length)	Format	Label	Valid Values	Comments
	CITY	C(24)			City 'U' = Unavailable ' ' = Unknown	
	STATE	C(2)			Postal code abbreviation for state 'U' = Unavailable ' ' = Unknown	
	ZIP	N(4)		Zip Code	5-digit zip code (nnnnn) .U = Unavailable . = Unknown	

Notes:

- 1a. The GEOCODE file is submitted for the *Youth* cohort only.
- 1b. Scramble Medical Record number different but linked to STUDYID in local table. ID may be any combination of alphanumerics so long as they are unique among study participants and remain the same for each participant from cycle to cycle. Do not include leading, embedded or trailing blanks or special characters.
- 2a. Each HMO will select two POLLDATEs: February 2000 and 2001 POLLDATEs. A record will be created containing the address of the youth (responsible subscriber) on each POLLDATE.
- 2b. Youth should have a record for each POLLDATE when they were active. If they have no POLLDATE because they enrolled and quit between two POLLDATEs, or quit before their first POLLDATE, a single record should be created using the closest POLLDATE.

General file description:

This file contains Youth cohort members' address on a given POLLDATE. An encrypted STUDYID (GEOID) is used on this file.

Attachment I

Non-Standard ICD-9 Codes

Vaccine Safety Datalink Project

NCK, NWK, SCK, KPC, and HP:

<u>ICD-9 Codes</u>	<u>Definition</u>
780.3F	Febrile Seizure

SCK:

<u>ICD-9 Codes</u>	<u>Definition</u>
X00002	Post Surgical Complications
X00003	MRI
X00005	HEP Lock, IV Medication, PIC Line
X00006	Transfusion
X00007	Hickman Cath.
X00008	CT Pre-OP
X00009	EEG
X00010	Renal Scan
X00014	Swelling
X00015	Incision Drainage
X00016	Other Continued Treatment
X00017	Other Pre-OP
X00018	Cardiac Monitoring
X00019	Deceased
X00098	Diagnosis Not Clear

Attachment II

Numeric Length Specifications

Vaccine Safety Datalink Project

The VSD data dictionary specifies variables on which calculations may be performed as numeric. In order to ensure accuracy, regardless of which platform a person/site is operating, the UNIX, DOS, or Windows platform numeric lengths have been designated.

Below is a summary of all numeric variables found in the VSD data dictionary and the minimum variable length based on the platform used. Files delivered to CDC as contract deliverables are required to use the numeric lengths listed under the UNIX/DOS/WINDOWS column.

VARIABLE	VSD FILE	UNIX/ DOS/ WINDOWS	MAINFRAME
SAS Dates	ALL	4	3
CHLDWGHT	HMOBIRTH	4	3
GESTAGE	HMOBIRTH	3	2
MC2000	CONSTANT	3	3
SAMPLE	CONSTANT	4	3
DURATION	INPT	3	3
ZIP	ADDRESS GEOCODE	4	4

Attachment III

ICD-9 Code Conventions

Vaccine Safety Datalink Project

The VSD study uses ICD-9 codes to represent diagnoses assigned to children during different health care encounters. There are different methods of assigning these codes but the following conventions should be used when submitting ICD-9 coded data to CDC.

1. **Variable type and length.** The ICD-9 variable should be a character variable with a length of six (6).

2. **First letter.** The first character should one of the following:

- A. One of the digits 0-9 followed by another digit (0-9)
- B. The letter 'V' followed by a digit (0-8)
- C. The letter 'E' followed by a digit (8 or 9)
- D. The letter 'M' followed by a digit (8 or 9)

At times, HMOs may have local codes to depict special conditions, e.g., X123. Any use of these codes should have accompanying written documentation explaining what these codes are (included in Attachment I).

3. **Decimals.** The decimal in numeric and V codes should not be included as part of the code. **It is implied that the decimal will occur after the following places:**

- A. Numeric codes -- after the 3rd digit
- B. 'V' codes -- after the 3rd character (V plus two digits, then decimal)
- C. 'E' codes -- after the 4th character (E plus three digits, then decimal)
- D. 'M' codes -- uses a slash after the 5th character (M plus four digits then **slash followed by a digit representing the behavior**)

4. **Alignment.** All codes are to be left justified.

5. **Leading zeroes.** Leading zeroes are added to numeric ICD-9 codes so that the implied decimal **would appear after the third position.**

- A.** Numeric codes that are less than 100. should have one or two leading zeroes added before the first digit so the decimal would occur after the third digit.
- 1) 123.1 will have no zeroes added => '1231'
 - 2) 23.5 will have one zero added => '0235'
 - 3) 3.5 will have two zeroes added => '0035'

- B.** No leading zeroes are added to 'V', 'E' or 'M' codes.

Note: Because E codes use three digits after the leading E, the implied decimal will appear one position to the right of the implied decimal of numeric and V codes.

- 6. Trailing zeroes.** Trailing zeroes should not be added after the implied decimal unless they are part of the code.

Example: 234. => '234' and not '2340' or '23400'

- 7. Non-standard ICD-9 codes.** Any code that is not found in the ICD-9 code book should be explained in writing anytime they are used. These usually take one of two forms:

- A.** Trailing codes used to explain a code, e.g., F added at the end of a seizure code to indicate that it is febrile; and
- B.** Local codes that are adopted by individual HMOs (included in Attachment I).

Attachment IV

Facility Codes

Vaccine Safety Datalink Project

KPC

<u>Name of Facility</u>	<u>Facility Number</u>	<u>Start Date</u>
Arapahoe	08	01JAN 1995
Aurora Centrepont	32	*01JULY 1995
Baseline	34	*01JULY 1998
Boulder	15	01JAN 1995
Broomfield	17	01JAN 1995
East	06	01JAN 1995
Englewood	14	01JAN 1995
Franklin	02	01JAN 1995
Hidden Lake	26	*01MAY1995
Lakewood	03	01JAN 1995
Longmont	19	01JAN 1995
Skyline 36		*01MAY1997
Smoky Hill	12	01JAN 1995
Southwest	09	01JAN 1995
Westminster	07	01JAN 1995
Wheat Ridge	10	01JAN 1995
Regional Highline Mental Health	01	
Mental Health West	22	
Mental Health Central	23	
Mental Health North	24	
Foothills Mental Health	35	

Kaiser Permanente of Colorado Affiliated Facilities

<u>Name of Facility</u>	<u>Facility Number</u>
Boulder Community Hospital 101	
The Children's Hospital	102
St. Joseph Hospital	103
Lutheran Hospital	104
Swedish Medical Center	105

*Kaiser Permanente facilities opened after Jan 1, 1995

NCK:

<u>Facility</u>	<u>Medical Center</u>
11	Oakland
12	Richmond
14	Hayward
16	Fremont
20	Petaluma
21	Vallejo
22	Napa
23	KFR
25	Livermore
31	San Francisco
32	French Hospital
35	S. San Francisco
36	Novato
37	Redwood City
38	Milpitas
39	San Rafael
40	Fairfield
41	Santa Clara
42	Mountain View
43	Sunnyvale
45	Campbell
47	Pleasanton
49	Manteca
50	Shadelands
51	Walnut Creek
53	Martinez Psych
54	Martinez
55	Antioch
58	Santa Rosa
59	Roseville
60	South Sacramento
61	Sacramento
62	Davis
63	Santa Teresa
67	Vacaville
68	Stockton
71	Rancho Cordova
72	Fresno
75	Gilroy
80	Oakhurst
83	Stanislaus
90	Out of Area
1A	Alta Bates

1B	Summit
1D	Mt. Diablo Hospital
99	Not Applicable

SCK:

<u>Facility</u> ¹	<u>Medical Center</u>
002	Sunset
052	Bellflower
101	Harbor City
151	Fontana
202	San Diego
251	Panorama City
305	West Los Angeles
354	Anaheim
401	Woodland Hills
501	Riverside

¹See SCK's Medical Care Program Area/Location Codes document for additional facility codes.

Attachment V

Startup Dates for NCK Facilities

Vaccine Safety Datalink Project

The following table shows the date that each facility began computerized immunization tracking for the VSD study. There is no date listed for Vacaville because it began tracking immunizations on the same day that it opened, and hence there is no potential problem of immunizations given before the tracking system was in place.

<u>Facility</u>	<u>Start Date</u>
Antioch	04NOV1991
Davis	08MAR1993
Fairfield	20MAY1993
Fremont	08MAR1993
Fresno	01OCT1992
Gilroy	20APR1993
Hayward	08DEC1992
Milpitas	01APR1993
Mountain View	01APR1993
Martinez	04DEC1991
Napa	21MAY1993
Novato	01JUN1993
Oakland	12FEB1991
Petaluma	26MAY1993
Pleasanton	28JUN1993
Richmond	01JUN1993
Rancho Cordova	18FEB1992
Roseville	18FEB1992
Redwood City	11MAR1991
Sacramento	21FEB1991
Santa Clara	13FEB1991
San Francisco	14FEB1991
Shadelands	16APR1992
San Rafael	01JUN1993
Santa Rosa	25MAY1993
South Sacramento	12MAR1991
So San Francisco	18MAR1991
Stockton	25JUN1993
Santa Teresa	12FEB1991
Vacaville	.
Vallejo	24MAY1993
Walnut Creek	13FEB1991